



November 22, 2024

Sent via ECF

Hon. Ona T. Wang
Daniel Patrick Moynihan
United States Courthouse
500 Pearl Street
New York, NY 10007-1312

Orrick, Herrington & Sutcliffe LLP
The Orrick Building
405 Howard Street
San Francisco, CA 94105-2669
+1 415 773 5700
orrick.com

Annette L. Hurst

E ahurst@orrick.com
D +1 415 773 4585
F +1 415 773 5759

RE: *The New York Times Co. v. Microsoft Corp., et al.* No. 1:23-cv-11195
New York Daily News, LP, et al. v. Microsoft Corp., et al., No. 1:24-cv-3285
The Center for Investigative Reporting, Inc. v. OpenAI, Inc., No. 1:24-cv-04872

Dear Judge Wang:

Per the Court's request during the October 30th status conference, on behalf of all parties in the above-referenced actions, counsel for Microsoft hereby submits the below updated chart containing a summary of discovery disputes for discussion at the upcoming December 3rd conference.

Given the number of issues included in the chart below, the parties have endeavored to list the issue categories in order of priority, alternating between Plaintiffs and Defendants.

Sincerely,

/s/ *Annette L. Hurst*

Annette L. Hurst

Key: Green represents Plaintiffs' filing; Purple represents Defendants' filing; Orange represents joint filing

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
I. Disputes Related to Plaintiffs' Requests for Inspection of OpenAI's Models							
305 ¹	11/1/24	305	OpenAI's Training Data	News Plaintiffs	OpenAI	Submitted joint dispute letter per Court's order (Dkt. 304).	<p>Plaintiff's position: The Times and the Daily News Plaintiffs request an order compelling OpenAI to identify and admit which of Plaintiffs' works were used to train each of the GPT-1, GPT-2, GPT-3, GPT-3.5, GPT-3.5 Turbo, GPT-4, GPT-4 Turbo, and GPT-4o families of models. Because this information is solely within Defendants' possession, News Plaintiffs have served numerous discovery requests since February seeking this information, including Requests for Production, Interrogatories, and Requests for Admission, and now request that OpenAI be ordered to provide this information instead of requiring Plaintiffs and their experts to spend weeks/months in OpenAI's tightly controlled "sandbox."</p> <p>As an update, the News Plaintiffs on November 20 filed a short letter to describe additional problems they have encountered in the "sandbox," including how OpenAI engineers erased a week's worth of data generated by the News Plaintiffs' experts during their inspection. Dkt. 328. Problems like these further support the News Plaintiffs' request for an order compelling OpenAI to identify the works on which its models were trained. Finally, OpenAI's suggestion below that the News Plaintiffs could have begun their inspections in June is not true. OpenAI did not make any training data available until October, <i>after</i> the parties reached agreement on an inspection protocol. See Dkt. 254.</p> <p>OpenAI's position: Plaintiffs recently served an effective total of almost 500 million requests for admission, and they have neither filed a motion to compel on these nor even <i>asked</i> to meet and confer. Indeed, Plaintiffs rushed to Court asking for relief before OpenAI's responses were even due. The information that Plaintiffs seek is in Plaintiffs' hands because OpenAI made training data available to them in early June (though Plaintiffs did not schedule their first inspection until October). And to accommodate Plaintiffs' requests to access several hundred terabytes of unstructured textual training data, OpenAI built an enterprise-grade virtual machine and installed hundreds of software tools for Plaintiffs' exclusive use. An order compelling OpenAI to respond to nearly 500 million requests for admission cannot be justified under Rule 26. Fortunately, a more reasonable approach exists. The</p>
328	11/20/24	345				News Plaintiffs filed a supplemental status update on November 20, 2024 (Dkt. 328).	

¹ Unless otherwise indicated, all docket numbers are from *New York Times Company v. Microsoft Corp., et al.*, Case No. 1:23-cv-11195-SHS.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							Court should direct Plaintiffs to collaborate with OpenAI to develop a plan for reasonable, targeted searches to be executed either by Plaintiffs or OpenAI.
250	9/27/24	250	OpenAI's Model Inspection Protocol	News Plaintiffs	OpenAI	Parties outlined their positions in September 27 status letter and await the Court's ruling	<p>Plaintiff's position: The Times and Daily News Plaintiffs request an order compelling OpenAI to provide Plaintiffs sufficient retail "credits" to analyze the extent to which unauthorized copies of Plaintiffs' copyrighted works are encoded within the parameters of the GPT LLMs themselves and the propensity of the GPT LLMs to disseminate or publicly display those works in outputs.</p> <p>OpenAI below proposes an "iterative approach." The News Plaintiffs have no objection to an iterative approach; however, OpenAI will only offer a very small amount of initial credits, and they want to charge the News Plaintiffs beyond those initial credits the same commercial prices that OpenAI charges its customers. In effect, OpenAI seeks to profit from this discovery. The News Plaintiffs have repeatedly asked OpenAI to disclose the actual costs they will incur from this inspection, but OpenAI has refused to provide that information.</p> <p>OpenAI's position: OpenAI has already agreed to provide Plaintiffs \$15,000 each in API credits, and to negotiate in good faith about any future credits that Plaintiffs may need with the expectation that the parties split the cost evenly. This iterative approach will provide the parties real world experience with running efficient searches, and also ensures that Plaintiffs are not paying "retail" pricing. Plaintiffs have not conducted any model inspections to date and have not used any of the \$15,000 API credits provided. Thus, their demand for \$800,000 in credits now is premature.</p>

II. Disputes Related to The Times's Search For and Production of Fair Use Documents

315	11/18/24	341	Factor Four Fair Use Evidence for Market Analyses	Microsoft	NYT	Microsoft and NYT have submitted letter briefing.	<p>Microsoft's position: Microsoft seeks the production of (1) information regarding subscription losses, including user surveys and other analyses of the reasons for subscription cancellation, which would directly refute The Times's claim that generative AI has caused subscription loss; (2) information regarding advertising revenue and analyses of effects on advertising revenue over time for the same purpose; (3) web traffic documentation to refute alleged effects on referral traffic; and (4) high level economic and financial reporting data regarding The Times's</p>
-----	----------	-----	---	-----------	-----	---	--

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							<p>business performance from 2015 to 2024, in order to model economic performance prior to and independent of the pandemic (RFP Nos. 8, 9, 75, 76, 77, 80, 89, and 101). The only way to assess whether Defendants' uses are a <i>substitute</i> is to assess the <i>effect, if any, upon the market for the works</i>; doing that requires a complete assessment of Plaintiff's business when that business is to monetize the works at issue.</p> <p>Plaintiff's position: The Times has already agreed to produce documents regarding the value of the asserted works and the market, including documents concerning any valuation of those works from 2015 to present, documents sufficient to identify the well-established market for The Times to provide paid access to and use of its copyrighted works, and documents sufficient to identify lost market value, among other relevant discovery set forth in Exhibit 1. This discovery is more than enough for Microsoft to address the fourth fair use factor, which focuses on whether Defendants' use of Times works provides a competing substitute in the market. Microsoft's expansive demands for every email regarding website traffic, subscription cancellations, or changes in its advertising revenue are overly broad and disproportionate to the needs of the case.</p>
310	11/18/24	333	OpenAI RFPs RFPs 141-148 (communications and economic analyses re market harm); RFPs 18, 33, 160-161 (communications and analyses re licensing of asserted works)	OpenAI	NYT	OpenAI and NYT have submitted letter briefing.	<p>OpenAI's position: OpenAI seeks production of evidence critical to its fair use defense. Specifically, OpenAI seeks an order compelling the Times to produce: (1) documents and communications regarding the Times's alleged market harm, i.e., decreases in revenue and traffic, and analyses of the causes; and (2) documents and communications regarding the markets for licensing the Times's works. These categories of documents go directly to the fourth factor of the fair use inquiry—"the effect of the use upon the potential market for or value of the copyrighted work," 17 U.S.C. § 107(4)—for which OpenAI bears the burden. The Times does not meaningfully contest the relevance of these categories of information, but instead makes unsubstantiated burden claims and incorrectly asserts that their responses to other RFPs obviate the need for full responses to the RFPs at issue.</p> <p>Plaintiff's position: The Times has already agreed to produce a significant number of documents concerning web traffic, revenue, market harm, subscriptions, and licensing. These documents are sufficient to meet OpenAI's legitimate discovery needs, and the additional discovery that OpenAI seeks is overbroad and unduly</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							burdensome. In addition to the voluminous discovery The Times already agreed to on the same topics, these requests seek granular data such as real-time monitoring of individual articles on The Times's website or daily internal communications regarding all licensing agreements that are irrelevant and disproportionate to the needs of this case. OpenAI's request is also a repeat of its pending October 22 motion (Dkt. 276) and should be denied for the same reasons (Dkt. 293).
276	10/22/24	293	OAI's RFPs 35 & 37 (Times' analyses of its trends and performance)	OpenAI	NYT	OpenAI and NYT have submitted letter briefs. The Court has not yet heard full argument and instructed the parties to address the issue in this chart.	<p>OpenAI's position: OpenAI seeks production of (1) documents and communications relating to formal or informal analyses of trends in readership or online subscriptions, in addition to the causes thereof; and (2) documents and communications relating to formal or informal analyses of the performance of content published on the Times' website, including analyses of methods to evaluate that performance. These materials relate directly to the harm alleged in the complaint, and they are important for OpenAI's fair use defense.</p> <p>Plaintiff's position: The Times agreed to produce (1) documents sufficient to show trends in readership and online subscriptions (RFP 35) and (2) documents sufficient to show trends in revenue from The Times's online content (RFP 37). Separately, The Times agreed to produce analyses regarding whether decreases in traffic to The Times's website were caused by Defendants' generative AI products or other market forces, as well as all documents and communications regarding the harm alleged in this case. This discovery is more than sufficient for OpenAI to assess The Times's damages allegations and the effect of OpenAI's infringement on the potential market for The Times's copyrighted works, which is the fourth fair use factor.</p>
276	10/22/24	293	OAI RFP 70 (Benefits of OpenAI's technology)	OpenAI	NYT	OpenAI and NYT have submitted letter briefs.	<p>OpenAI's position: OpenAI seeks production of documents and communications concerning the benefits of OpenAI's technology to the journalism industry. These materials relate to OpenAI's fair use defense and demonstrate the public benefits of OpenAI's technology.</p> <p>Plaintiff's position: The Court should deny this portion of OpenAI's motion for the reasons outlined in its November 22 Order denying OpenAI's Dkt. 236 motion to compel. <i>See</i> Dkt. 244 at 3-4. The "public benefits" inquiry of the fourth fair use factor looks at the public benefits of Defendants' use of Times works. The Times already agreed to</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							produce documents concerning Defendants' use of Times works, which will include documents about any purported "benefits" of Defendants' use, if any exist. Anything else OpenAI seeks in this discovery request is irrelevant and overbroad. As this Court held, the fair use defense does not require "discovery about Plaintiff's views on or statements about the 'public benefits' of Gen AI in journalism" more generally. Dkt. 244 at 4. "Whether nonparties' Gen AI tools confer benefits on the journalism industry is not relevant to a determination of whether Defendant's acts—i.e., the alleged copying involving Defendant's Gen AI tools—constitute fair use." <i>Id.</i> at 5.
276	10/22/24	293	OAI RFP 76 & 82 (impact of AI and use of AI on the Times)	OpenAI	NYT	OpenAI and NYT have submitted letter briefing.	<p>OpenAI's position: OpenAI seeks production of (1) documents regarding the impact of third-party generative AI products on journalism (that is, products offered by companies other than OpenAI), and (2) the Times' own use of AI products in support of its business. This topic overlaps to some extent with OpenAI's prior motion to compel (Docket No. 263), which is under submission. As with the issues raised in that prior motion, the requested documents are relevant to OpenAI's fair use defense.</p> <p>Plaintiff's position: The Court should deny this portion of OpenAI's motion to compel discovery for the same reasons it denied OpenAI's similar motion at Dkt. 236. <i>See</i> Dkt. 244 (Court's Order denying OpenAI's motion to compel discovery into The Times's use of generative AI and discovery regarding third party generative AI). Discovery into (1) the impact of third-party generative AI tools on the journalism market and (2) The Times's use of third-party generative AI tools for marketing is not relevant to OpenAI's fair use defense. The fair use inquiry is context-specific and is centered on <i>Defendants'</i> use of The Times's proprietary works and any public benefits that use may produce, a topic on which The Times has already agreed to produce documents. <i>See</i> Dkt. 244 at 2 ("Each of these [fair use] factors requires a scrutiny of a <u>defendant's</u> purported use of the copyrighted work(s), and whether that <u>defendant's</u> use may constitute 'fair use' under the Act.").</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
III. Disputes Related to Custodians and Custodial Documents							
316-Sealed ² 317-Public	11/18/24	330	OpenAI's request that the Times add 5 additional ESI custodians.	OpenAI	NYT	OpenAI and NYT have submitted letter briefing.	<p>OpenAI's position: OpenAI seeks the addition of the following five Times custodians: (1) David Rubin, Chief Brand and Communications Officer and Publisher of Wirecutter; (2) Joy Robins, the Times's Global Chief Advertising Officer; (3) Michael Greenspon, the Times' Global Head of Licensing and Print Innovation; (4) Hannah Masuga, former Technical Product Director of Machine Learning Platform; and (5) Sam Dolnick, Deputy Managing Editor. These custodians are likely to have unique, responsive information related to factor four of the fair use inquiry, the Times's allegations of harm, and the Times's exclusion protocols (including paywalls and bot blocking). If the Times added these five custodians, that would bring their total custodians from 18 to 23; by contrast the Times is seeking 38 OpenAI custodians, 12 more than the 26 custodians OpenAI has already designated.</p> <p>Plaintiff's position: The Times agreed to produce documents from Joy Robins, so that request is now moot. The Times opposes OpenAI's request to compel the addition of four new, duplicative custodians because The Times's 22 existing custodians are better positioned to address each of the topics OpenAI invokes as the basis of its motion. The Times has already designated numerous high-ranking custodians with relevant knowledge on: (1) the harm to The Times's brand, reputation, and trademarks, (2) licensing and the market for Times works, (3) the operation and enforcement of The Times's paywall and exclusion protocols, (4) and how The Times's journalism has been impacted by Defendants' conduct. The four additional custodians OpenAI seeks are not likely to yield relevant, non-duplicative discovery on any of these issues.</p>
309	11/18/24	329-OpenAI 331-MSFT	Defendants' preservation of, search for, and production of text messages and social media messages (The News Plaintiffs)	News Plaintiffs	OpenAI & Microsoft	News Plaintiffs and Defendants have submitted letter briefs, on which the Court has not	<p>Plaintiffs' Position: The News Plaintiffs request an order to establish a framework to govern all parties' collection and production of text messages and social media messages. This dispute relates to a similar (though not identical) motion filed in the <i>Authors Guild</i> case, Dkt. 231, which was argued during the October 30 status conference.</p> <p>With respect to social media messages, the News Plaintiffs incorporate the arguments raised by the <i>Authors Guild</i> plaintiffs, and the Court's</p>

² The Times does not intend to move to seal any of the exhibits attached to the motion at Dkt. 316.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
			request that OpenAI and Microsoft preserve, search for, and produce text messages and social media messages from all custodians)			yet heard argument.	<p>ruling on that dispute will likely apply to this case. With respect to text messages, since Plaintiffs filed their motion, OpenAI has now agreed to search for and collect text messages from all 26 agreed custodians, instead of just a select few custodians. That aspect of the dispute appears to have been mooted. But there are still two disputes with respect to Microsoft. First, the parties have a dispute about how responsive text messages should be produced. The News Plaintiffs propose that text messages be produced in 24-hour increments (i.e., each 24-hour exchange of text messages is produced as one document with non-responsive personal information such as medical information or communications about minor children redacted), which is important for capturing relevant context. Microsoft has not made a counter-proposal. Second, the News Plaintiffs request that Microsoft search through and produce documents from Slack channels that Microsoft employees used to communicate with OpenAI employees. Microsoft claims that only OpenAI can search these channels, but has not clarified whether OpenAI will search them for all of the agreed Microsoft custodians (as opposed to the agreed OpenAI custodians). If a Microsoft custodian used this OpenAI Slack channel to communicate with OpenAI employees who are not custodians, those communications are likely being missed.</p> <p>OpenAI's position: OpenAI has already agreed to conduct a reasonable investigation regarding its custodians' use of text messages. The News Plaintiffs contend that OpenAI has limited its investigation to a subset of its custodians. That is false, as the meet-and-confer correspondence attached to Plaintiffs' own motion makes clear. In it, OpenAI told Plaintiffs it "is collecting and producing documents from its agreed custodians . . . that the custodians used to communicate about the issues relevant to this action," including texts. Plaintiffs' characterization above as to OpenAI's agreement regarding text messages is incorrect; OpenAI's position is explained in detail at Dkt. 329. As to social media discovery, the News Plaintiffs fail to acknowledge that OpenAI has proposed a path forward that would get Plaintiffs the documents they want. OpenAI believes that California Labor Code § 980 prohibits OpenAI from asking its employees for access to their social media accounts. However, OpenAI has informed Plaintiffs that if they choose to serve subpoenas targeting OpenAI's employees' social media messages, OpenAI will cooperate with Plaintiffs and its employees to move that process</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							<p>forward efficiently. Instead of engaging with OpenAI on a logical, efficient, and reasonable process, the News Plaintiffs filed this motion. It should be denied.</p> <p>Microsoft's position: Microsoft already collected and processed text messages in an easy-to-review and usable format. Microsoft also confirmed that OpenAI is collecting and producing Slack messages involving Microsoft employees, since Microsoft has no ability to extract data from OpenAI's enterprise Slack accounts. And OAI itself has stated in its response brief that it "is already collecting and producing relevant and responsive documents from Slack channels controlled by OpenAI and that may have been used for communication with Microsoft employees." Dkt. 329 at n.1. Unsatisfied, News Plaintiffs insist that Microsoft pay to re-process the text message data into a different format and that it collect data that it cannot collect and that OpenAI is already collecting. There is nothing here for the Court to do but deny the motion.</p>
204 313	8/12/24 11/18/24	211 342	OpenAI's custodians (The Times seeks eight additional OpenAI custodians)	NYT	OpenAI	Addressed at the October 30th conference; OpenAI ordered to provide hit counts for the eight disputed custodians by November 13 (Dkt. 304) The News Plaintiffs on November 18 filed a letter brief (Dkt. 313) seeking four additional custodians and reiterating	<p>Plaintiff's position: The News Plaintiffs stand by their request for the following seven additional OpenAI custodians, who remain disputed: James Betker, Daniel, Kokotajlo, Brad Lightcap, Ryan Lowe, William Saunders, Ilya Sutskever, and Wojciech Zaremba. <i>See</i> Dkt. 204. OpenAI has now agreed to add William Fedus as a custodian. The court-ordered hit counts for these eight custodians, which OpenAI recently provided (Dkt. 308) further demonstrate why they should all be added. Based on those hit counts, OpenAI would produce around 25,000 additional documents from these custodians' files, which is proportional to the needs of these consolidated news cases. The News Plaintiffs' November 18 letter motion (Dkt. 313) cites additional recently produced documents to support these custodians' inclusion.</p> <p>Based on their ongoing review of Defendants' document productions and other information, the News Plaintiffs have also moved to compel four additional OpenAI custodians: Michael Lampe, Suchir Balaji, Peter Deng, and Ethan Stock. <i>See</i> Dkt. 313.</p> <p>In total, the News Plaintiffs now seek eleven additional OpenAI custodians and request a ruling on these custodians so the parties can promptly move forward toward depositions.</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
						their request for the previously briefed eight custodians.	<p>OpenAI's position: OpenAI has compiled hit counts pursuant to the Court's order. These figures show that adding the custodians Plaintiffs demand would require reviewing almost 123,000 documents after de-duplication, when including families. The fact that Plaintiffs apparently expect that onerous review to result in the production of a mere 25,000 documents does not, as they assert, establish that the search is proportionate to the needs of the case—it shows precisely the opposite. The low expected responsiveness demonstrates that Plaintiffs' search terms are overbroad and that they are attempting to force OpenAI to conduct a costly review without consideration for proportionality or burden. Although Plaintiffs suggest otherwise, the relevant question regarding proportionality is how many documents OpenAI will have to collect and review—not solely how many documents it will produce. It is not proportionate to the needs of the case to require OpenAI to review 123,000 more documents when it is already committed to reviewing half a million.</p> <p>As Plaintiffs complete their review of documents from OpenAI's 26 agreed-upon custodians, the parties will continue to meet and confer about any alleged gaps in OpenAI's productions, including the need for any additional custodians. To date, however, Plaintiffs have failed to identify any compelling reason why the currently designated custodians are inadequate or what relevant, non-cumulative information the requested custodians could provide that would justify the burden of collecting and searching their files. Nevertheless, in the spirit of compromise, OpenAI has offered substantial compromises to produce custodial information from four additional custodians (Lampe, Lightcap, Balaji, Fedus) and Plaintiffs should be ordered to accept these compromises.</p>
228	8/27/24	233	OpenAI's search terms (The Times requests that OpenAI run additional search terms for its custodians' ESI)	NYT	OpenAI	Addressed at the October 30th conference; the Court ruled that the total hit count will be capped at or around	<p>Plaintiff's position: Consistent with the Court's Order (Dkt. 304), The News Plaintiffs after the October 30 hearing prepared a revised search term proposal to OpenAI, which included additional terms for the newly consolidated CIR plaintiff. In part because of those new terms, that search term proposal initially yielded an increase in total hit counts, so the News Plaintiffs on November 21 submitted a revised and narrowed proposal, aimed at generating at or around 500,000 total hits for the 26 agreed custodians. OpenAI's response to this second proposal is due</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
						500,000 for the current set of agreed custodians (Dkt. 304) and the parties will meet and confer about the precise search terms and rolling productions.	<p>before the December 3 conference, and the News Plaintiffs will be prepared to discuss the parties' progress at the hearing.</p> <p>OpenAI's position: Plaintiffs sent OpenAI revised search term proposal on November 13. Those terms yielded 659,407 hits when run across the existing 26 custodians—an increase from the 632,357 hits yielded by Plaintiffs' prior proposal. Consistent with the Court's October 31 Order (Dkt. 304), which set a 500,000-hit cap on Plaintiff's search terms, OpenAI proposed a counterproposal that yielded 335,919 hits—an increase from the 195,362 hits yielded by its prior proposal. OpenAI offered to review these documents and revisit search terms at a later date, but Plaintiffs circulated another search term proposal on November 21, the day before this submission. In accordance with the Court's guidance and the terms of the ESI Order (Dkt. 251), OpenAI is currently analyzing that search term proposal and will respond to Plaintiffs next week. OpenAI continues to review documents hitting on the parties' agreed upon terms.</p>

IV. Disputes Related to The Times's Search for and Production of Fair Use & Substantial Non-Infringing Uses Evidence

320-Sealed ³ 321-Public (319-Motion to Seal)	11/18/24	332	The Times's Development and Use of GenAI (Fair Use & Substantial Non-Infringing Uses)	Microsoft	NYT	Microsoft and NYT have submitted letter briefing	<p>Microsoft's position: Microsoft seeks the production of (1) documents reflecting efforts by The Times to train, fine-tune, or otherwise develop and use a Generative AI tool (RFP Nos. 31, 32, 102, and 106); and (2) a survey conducted by The Times concerning workplace generative AI usage (RFP No. 108). These are not the same as the documents requested by OpenAI that were addressed in ECF 344, and that Order does not address or resolve Microsoft's arguments for relevance. The Times's efforts to train its own AI are directly relevant under <i>Google v. Oracle</i>, where plaintiff's failed efforts to develop the technology at issue weighed against it on the fourth factor. Further, the Court did not address the fact that Plaintiff's uses of the tools that result from the copying are directly relevant to assessment not only of the transformative nature of the tools but also to the completely separate defense of substantial non-infringing uses.</p> <p>Plaintiff's position: The Court should deny Microsoft's motion for the same reasons that it denied OpenAI's motion to compel discovery into</p>
---	----------	-----	---	-----------	-----	--	--

³ The Times does not intend to move to seal any portions of the letter brief or any exhibits attached to the motion at Dkt. 320.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							The Times's use and creation of its own and third-party generative AI at Dkt. 236. <i>See</i> Dkt. 344. Microsoft's motion is duplicative of OpenAI's two prior motions to compel the same discovery. <i>See</i> Dkt. 236 (denied), Dkt. 276 (pending). Relevance issues aside, The Times already agreed to discovery into The Times's use of Defendants' generative AI tools. Discovery into The Times's development and training of its own generative AI tools and its employees' use of generative AI is irrelevant to fair use because factors one and four are about <i>Defendants'</i> use of Times works—not The Times's use of its own generative AI or its own works. As this Court held, “whether the copyright holder has used tools in the defendant’s general industry” is not relevant to any of the fair use factors. Dkt. 344 at 3. Discovery into The Times's use of generative AI or its use of its own works to train generative AI is also irrelevant to Microsoft's “substantial non-infringing use” defense, which examines the infringing uses of Defendants' generative AI products by others. The Times cannot infringe its own works.

V. Disputes Related to Depositions

306	11/1/24	306	Early 30(b)(6) depositions	All Parties	All Parties	Submitted joint dispute letter per Court's order (Dkt. 304)	<p>Plaintiff's position: The News Plaintiffs seek two early 30(b)(6) custodial depositions prior to the upcoming December 3, 2024 discovery conference. Specifically, Plaintiffs seek one deposition of OpenAI and one deposition of Microsoft, each limited to 7 hours, and each shared with the Class Plaintiffs in the <i>Authors Guild</i> case. These depositions will “help focus document discovery” and resolve many of the parties’ disputes. <i>See</i> 10/30/24 Tr. 8:2-20. OpenAI's assertion below that Plaintiffs are “refusing” to produce their own 30(b)(6) custodial witness is incorrect.</p> <p>Microsoft's position: There is no basis for a custodial deposition of Microsoft. Plaintiffs have raised no issue with the court relating to Microsoft's custodians, sufficiency of document collection, or document retention, so there is no basis for discovery about discovery from Microsoft. To the extent Plaintiffs seek a deposition relating to output data, that is not a “custodial” deposition; it is a merits deposition. And Plaintiffs did not previously request it; they have no grounds for taking it on short notice before the next conference; and such an accelerated schedule would pose a substantial burden. Additionally, this issue was never fully briefed as Microsoft had ¾ of a page to address its position in the joint letter filed with the Court. Nor did Plaintiffs serve any notice, nor seek to meet and confer about any of their proposed</p>
-----	---------	-----	----------------------------	-------------	-------------	---	--

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							<p>deposition topics before the Status Conference. The Court should deny Plaintiffs' requested relief as to Microsoft.</p> <p>OpenAI's position: OpenAI is conferring with Plaintiffs regarding early custodial 30(b)(6) depositions, but Plaintiffs are seeking a deposition from OpenAI on overbroad, privileged and improper topics, while refusing to provide any custodial 30(b)(6) themselves. The parties' early custodial 30(b)(6) depositions should be mutual, cabined to 7 truly custodial topics per side, limited to 3-hours each, and take place in January 2025.</p>
311	11/18/24	338	30(b)(6) Deposition Regarding Microsoft's Log Data	News Plaintiffs	Microsoft	News Plaintiffs and Microsoft have submitted letter briefs, on which the Court has not yet heard argument.	<p>Plaintiffs' position: The News Plaintiffs request an order compelling Microsoft to provide a witness for a 30(b)(6) deposition prior to December 20, 2024, on the topic of Microsoft's log data reflecting user sessions and associated output logs, without prejudice to take a subsequent 30(b)(6) deposition. This data, including the generative output provided to end-users, is important to News Plaintiffs' claims because it not only serves as evidence of copyright infringement, but also informs the damages and fair use analysis as to whether Defendants' products are competing as substitutes for the News Plaintiffs' copyrighted works. A deposition on the log data now will help streamline the issues in the case because of the opacity of the sample log data Microsoft has produced thus far and Microsoft's representations that the full log data is extremely voluminous and burdensome to search. This issue is related to the issues briefed in the parties' joint letter filed on November 1 concerning taking early 30(b)(6) custodial depositions of each Defendant (Dkt. 306), and the topics concerning logs overlap in substance with topics 15 and 16 in Exhibit 1 attached to that letter (Dkt. 306-1).</p> <p>Microsoft's position: Microsoft is retaining more than 15 petabytes of user log data across five logging formats comprising more than 300 separate fields, much of which is totally irrelevant. The data cannot be reproduced and can only be inspected through querying. Microsoft voluntarily produced documentation and created samples of the data to help Plaintiffs understand it. Rather than ask a single question about the documentation and samples, Plaintiffs short circuited the meet and confer process after the October 30 hearing in the guise of seeking an immediate "custodial" deposition. This Court's decision in <i>Winfield v. City of New York</i>, No. 15-cv-05236, 2018 WL 840085, at *3 (S.D.N.Y.</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							Feb. 12, 2018) compels denial of Plaintiffs' motion. As the Court did in <i>Winfield</i> , it should instruct Plaintiffs to first examine the data samples, discuss the samples with their experts, and seek additional information informally from Microsoft in order to craft relevant and sufficiently specific deposition topics. Any deposition should be conducted no sooner than 30 days after such a notice is served.
261	10/23/24	278	Deposition Protocol	All Parties	All Parties	News Plaintiffs and Defendants submitted simultaneous letter briefs. At the October 30 hearing the Court denied the relief requested as premature.	Parties' Joint position: The parties understand the Court's view from the October 30 Status Conference that discussion of a deposition protocol is premature given that the parties are continuing to negotiate and raise disagreements relating to custodians and document collection and production. However, the parties do believe that it is important that a deposition protocol be ordered in these cases. To move that process forward, the parties intend to continue the meet and confer process and submit renewed briefing on this issue in advance of the next status conference in early 2025. The parties respectfully request that this issue be taken up by the Court at that time to ensure that a process is in place before depositions begin.

VI. Disputes Related to the Scope of Products Subject to Discovery

279	10/23/24	294	Scope of Microsoft's products subject to discovery	News Plaintiffs	Microsoft	NYT and Microsoft have submitted letter briefs, which the Court has not yet heard argument on	<p>Plaintiff's position: The News Plaintiffs request an order compelling Microsoft to search for and produce 1) documents responsive to The Times's RFP 67 regarding its plans to commercialize existing generative AI products built with Times content, including but not limited to Bing Chat and Microsoft Copilot; and 2) documents concerning Copilot Daily—Microsoft's recently announced AI-powered news product—and any other generative AI products that Microsoft is currently developing in categories that compete with News Plaintiffs' products—namely news, cooking, consumer reviews, and sports. News Plaintiffs allege an ongoing and continuous course of illegal conduct, which entitles them to discovery on Microsoft's post-complaint conduct, including with respect to products that remain in development.</p> <p>Microsoft's position: Microsoft requests a protective order that the scope of discovery encompasses only those products released to the public at the time Plaintiffs' respective complaints were filed and actually identified in the complaints. Discovery into products released after the lawsuits were filed, and into new products that may or may not even be released, should be denied. In particular, discovery regarding</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							Copilot Daily should be denied, as that product is based on an entirely new technology stack and business model not named in the complaint, and that was only released a few weeks ago in October.
287	10/24/24	302	Scope of OpenAI's products subject to discovery	News Plaintiffs	OpenAI	NYT and OpenAI have submitted letter briefs, which the Court has not yet heard argument on	<p>Plaintiff's position: The News Plaintiffs in this motion sought an order compelling OpenAI to search for and produce documents concerning SearchGPT (a GPT-based search product powered by News Plaintiffs' content), the OpenAI API, Copilot, and any other current or anticipated GPT-based product (including CustomGPTs) that relate to content in categories that compete with News Plaintiffs' products—namely news, cooking, consumer reviews, and sports. News Plaintiffs allege an ongoing and continuous course of illegal conduct, which entitles them to discovery on OpenAI's post-complaint conduct, including with respect to anticipated GPT-based products.</p> <p>As an update, after The News Plaintiffs filed their motion, OpenAI agreed to provide discovery about SearchGPT, the OpenAI API, and Copilot. Dkt. 302. The dispute has therefore been narrowed to whether OpenAI should also provide discovery into anticipated GPT-based products (including CustomGPTs) that relate to content in categories that compete with News Plaintiffs' products.</p> <p>OpenAI's position: OpenAI has already agreed to produce documents about every specific product Plaintiffs have identified, i.e., ChatGPT (including Browse (f/k/a Browse with Bing)), SearchGPT, the OpenAI API, and Microsoft's Copilot. The scope of Plaintiffs' request for documents concerning "any other current or anticipated GPT-based product (including CustomGPTs)" is overly broad and vague, as it encompasses a sweeping, undefined and unknown range of third-party products and other product categories. OpenAI remains committed to continuing to meet and confer regarding Plaintiffs' requests.</p>

VII. Disputes Related to OpenAI's Alleged Acquisition of Data

322-Sealed 323-Public	11/19/24	339	The Times's RFP 49 (documents concerning Defendants' knowledge of whether Defendants'	NYT	OpenAI	The Times and OpenAI have submitted letter briefs, on which the Court has not	<p>Plaintiff's position: The Times requests an order requiring OpenAI to produce documents "concerning Defendants' monitoring, tracking, and knowledge of investigations into whether Defendants' AI Models contain copyrighted content, including Times Content." This request is relevant to The Times's allegation of willful infringement. FAC ¶¶ 124-26. The dispute has been narrowed to whether OpenAI should limit productions to documents that specifically address The Times, the Daily</p>
--------------------------	----------	-----	---	-----	--------	---	---

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
			models contain copyrighted content)			yet heard argument.	<p>News Plaintiffs, and the CIR. The News Plaintiffs appropriately seek a broader scope, which should include documents addressing OpenAI's awareness that it was using copyrighted journalism content more generally.</p> <p>OpenAI's position: The only dispute relates to The Times's expansive definition of "journalism." It should be rejected as imposing an enormous burden relating to irrelevant information.</p>
322-Sealed 323-Public	11/19/24	339	The Times's RFP 95 (documents concerning Defendants' "Data Working Group")	NYT	OpenAI	The Times and OpenAI have submitted letter briefs, on which the Court has not yet heard argument.	<p>Plaintiff's position: The Times seeks an order compelling OpenAI to produce documents concerning Defendants' "Data Working Group," which OpenAI and Microsoft formed in early 2023 to align their approaches to acquiring content for model training datasets. These documents are relevant to numerous allegations, including how Defendants prioritize high-quality content for model training. OpenAI has refused to search for responsive documents, claiming that "all of the materials sought" are "certain to be privileged." That assertion is untenable, including because the documents will include communications between OpenAI and Microsoft employees, including non-legal employees. Moreover, OpenAI has provided no information about the number of documents at issue, undermining their claim that it would be unduly burdensome to conduct a privilege review.</p> <p>OpenAI's position: The Times is seeking discovery of privileged materials and forcing OpenAI to collect files from its lawyers and engage in a logging exercise is disproportionate and unduly burdensome.</p>
283	10/24/24	301	The Times's RFP 26 (documents concerning OpenAI's use of web crawlers)	NYT	OpenAI	NYT and OpenAI have submitted letter briefs, which the Court has not yet heard argument on.	<p>Plaintiff's position: The Times requests an order compelling OpenAI to produce documents concerning (i) Microsoft web crawlers and (ii) web crawlers used by developers of customGPTs. For (i), The Times alleges that Microsoft provides these web crawlers to assist OpenAI with accessing journalism content, including Times content. FAC ¶ 95. Documents produced in discovery substantiate that allegation. See Mot. Ex. H (Dkt. 283-8) (Microsoft-OpenAI agreement concerning Microsoft's provision of "web crawling services" to OpenAI). For (ii), the News Plaintiffs allege that OpenAI's Custom GPT Store contains Custom GPTs specifically designed to circumvent publishers' paywalls, including a "'Remove Paywall' Custom GPT" and a "'News Summarizer' Custom GPT." Daily News Compl. ¶ 147. The requested</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							<p>documents will reveal how OpenAI designed technologies to enable others to circumvent paywalls and harm publishers like News Plaintiffs.</p> <p>OpenAI's position: OpenAI already agreed to conduct a reasonable search for and produce documents concerning OpenAI's use of web crawlers to crawl content of articles published by The Times in connection with training the relevant models. OpenAI has no control over how Microsoft or third-party GPT developers use web crawlers. Nonetheless, OpenAI remains committed to meeting and conferring about this RFP.</p>
322-Sealed 323-Public	11/19/24	339	The Times's RFP 17 (documents concerning the compilation of Defendants' Training Datasets)	NYT	OpenAI	Resolved.	<p>Parties' Statement: OpenAI has agreed to conduct a reasonable search for relevant, non-privileged documents in its possession, custody, and control, if any, discussing the compilation, acquisition, and curation of the text training data used to train the relevant models. To the extent relevant, non-privileged documents in OpenAI's possession, custody, or control discussing OpenAI's approach to paywalls with respect to training data—including the impact of paywalls with respect to training data—for the relevant models and products exist and can be located pursuant to a reasonable search, those documents will be produced.</p>
322-Sealed 323-Public	11/19/24	339	The Times's RFP 24 (documents concerning analyses of the value of including Journalism content in Defendants' Training Datasets)	NYT	OpenAI	Resolved.	<p>Parties' Statement: OpenAI has agreed to conduct a reasonable search for relevant, non-privileged documents in its possession, custody, and control, if any, discussing any evaluation or assessment conducted, generated, or commissioned by OpenAI relating to the value of including Journalism content in the text training datasets used to train the relevant models. OpenAI also agrees to include within its scope responsive documents discussing evaluations or assessments prepared by entities other than OpenAI. To the extent such documents are within OpenAI's possession, custody, or control, OpenAI will (a) conduct a reasonable investigation to locate responsive documents and (b) conduct a custodial search based on search terms to be negotiated with The Times.</p>
VIII. Disputes Related to OpenAI's Alleged Use of Content							
322-Sealed 323-Public	11/19/24	339	The Times's RFPs 86-87 (documents concerning OpenAI's creation and use)	NYT	OpenAI	The Times and OpenAI have submitted letter briefs, on which the Court has not	<p>Plaintiff's position: The Times seeks an order requiring OpenAI to produce documents concerning OpenAI's creation, use, and contemplated use of synthetic datasets. "Synthetic" data refers to "artificial, algorithmically-manufactured data, including data generated using a generative AI model." Dkt. 283-2 at 4. Put differently, OpenAI uses models trained on Times content and other journalism content to</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
			of synthetic datasets)			yet heard argument.	<p>create “synthetic” datasets that mimic Times works. OpenAI’s creation and use of synthetic datasets is relevant to fair use because these datasets may impact the market for training data.</p> <p>With respect to RFP 86, which addresses OpenAI’s creation of synthetic datasets, The Times seeks documents concerning the use of journalism content to create synthetic datasets that mimic and ultimately substitute for authentic journalism. OpenAI seeks to improperly limit productions to documents concerning the use of Times content for synthetic data creation. With respect to RFP 87, which addresses OpenAI’s use or contemplated use of synthetic datasets, OpenAI refuses to produce documents concerning contemplated uses, which would exclude, among other relevant material, documents concerning how OpenAI tried (unsuccessfully) to use synthetic data. These attempted uses are relevant to demonstrate the value of authentic journalism content for training. Finally, for both RFPs 86 and 87, OpenAI seeks to limit productions to a “sufficient to show” production.</p> <p>OpenAI’s position: RFP 86 asks for “All Documents concerning use of Generative AI Models to create Synthetic Datasets.” The Times narrows the focus in its letter-brief to The Times’s content. OpenAI agrees that tying this request to Plaintiffs’ copyrighted content is an appropriate scope, and has therefore agreed to produce documents sufficient to show the use of OpenAI models, if any, trained on Times Content to create synthetic datasets. But The Times now takes issue with the “sufficient to show” limitation. That language is meant to reflect Plaintiffs’ narrowing and the parties’ compromise, and thus prevent an unbounded inquiry into the scope of RFP 86 as written – i.e., the development of every piece of synthetic data ever created on any subject and for any purpose.</p> <p>RFP 87 asks for “All Documents concerning [OpenAI’s] use or contemplated use of Synthetic Datasets, including to train Your Generative AI Models.” The Times’s letter-brief focuses its request to the market or substitution value (either now or in the future) of OpenAI’s synthetic data, as compared to “authentic journalism content” from The Times or other textual news sources. OpenAI is willing to produce documents sufficient to show value assessments for synthetic data as compared to written news. OpenAI’s willingness to do so stands in stark contrast to The Times’s continual refusal to produce market-</p>

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							<p>related discovery. Like for RFP 86, OpenAI's "sufficient to show" language is not included to withhold market or value documents related to written news, but to focus Plaintiffs' unbounded request for "all documents concerning [OpenAI's] use or contemplated use of synthetic datasets" to what The Times agrees is the narrower scope of this case. OpenAI uses—and contemplates using—synthetic data for purposes that do not relate to textual news and are not substitutive of anything. Accordingly, for both RFPs 86 and 87, OpenAI requests the Court accept OpenAI's "sufficient to show" limitation aimed at the core of what The Times seeks, and deny The Times's motion.</p>
322	11/19/24	339	The Times's RFP 2 (documents OpenAI has submitted to governmental entities relating to the allegations in the Complaint)	NYT	OpenAI	The Times and OpenAI have submitted letter briefs, on which the Court has not yet heard argument.	<p>Plaintiff's position: The Times requests an order requiring OpenAI to (i) identify any relevant proceedings in which it has produced documents related to the issues in this case, which will facilitate additional negotiations about the appropriate scope of OpenAI's response to RFP 2; and (ii) produce the Interrogatory responses numbered 2, 3, 5, 6, 7, and 10 that OpenAI submitted to the FTC for Civil Investigative Demand No. 232-3044, which are likely to contain information relevant to this case. To resolve a dispute in the Authors Guild case, OpenAI previously produced responses to some of these FTC Interrogatories, but not the ones requested in this motion. <i>See Authors Guild</i>, Case No. 23-cv-8292, Dkt. 106.</p> <p>OpenAI's position: OpenAI has agreed to produce the final FTC Interrogatory responses, and without conceding that this request seeks relevant or proportional documents, there are no other U.S. government proceedings where OpenAI has produced documents "that concern or relate to the allegations in the Complaint."</p>
283	10/24/24	301	The Times's RFPs 84–85 (documents concerning OpenAI's knowledge and use of Dnyuz and similar websites that copy publisher content)	NYT	OpenAI	Resolved.	Parties' position: Since the November 6, 2024 chart submission, the parties have resolved this dispute. For RFP 84, OpenAI has now agreed to run "Dnyuz" as a search term and to produce responsive documents. Dkt. 307. For RFP 85, OpenAI has now agreed to search for and produce documents concerning three exemplar sites identified by The Times, and to use search terms to capture documents responsive to the RFP.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
			without permission)				

IX. Disputes Related to Licensing and Data Access Documents

141	6/11/24	147	The Times's RFP 6 (communications regarding licensing and data agreements)	NYT	OpenAI	NYT and OpenAI have submitted letter briefs, which the Court heard argument on at the September 12 conference; parties ordered to further meet and confer, and the parties continue to do so.	<p>Plaintiff's position: The Times seeks an order requiring OpenAI to produce internal and external communications relating to agreements in which it licenses content to use with its generative AI models and products.</p> <p>The parties are close to resolving this dispute because OpenAI has now agreed to produce internal and external communications relating to its negotiation of licensing agreements. Dkt. 307 at 7.</p> <p>The sole remaining issue is whether these productions should be limited to the currently agreed upon 26 custodians, or whether OpenAI should also produce documents from other employees involved in these licensing negotiations. As an update, the News Plaintiffs have proposed that OpenAI produce licensing-related communications from two additional employees who worked on these deals: Emily Stern and David Duxin. To be clear, these two employees would only be custodians for purposes of this RFP, and the search terms for them would be limited to search terms focused on the licensing issue.</p> <p>OpenAI's position: OpenAI has produced all relevant data access agreements for which third party confidentiality concerns have been resolved. OpenAI has also agreed to produce non-privileged relevant communications between (a) OpenAI and third-parties related to data access agreement negotiations, and (b) between OpenAI employees related to data access agreement negotiations. OpenAI is committed to continuing to meet and confer about any alleged gaps in OpenAI's productions, including the need for additional custodians, after The Times completes its review of such documents.</p>
-----	---------	-----	--	-----	--------	---	---

X. Disputes Related to Documents Concerning OpenAI's Alleged Transition to a For-Profit Company

141	6/11/24	147	The Times's RFP 11 (OpenAI's alleged transition to a for-profit company)	NYT	OpenAI	Resolved.	<p>Parties' position: Since the November 6, 2024 chart submission, the parties have resolved this dispute. OpenAI has agreed to conduct a reasonable search of its designated custodians and produce documents relating to OpenAI's creation of a for-profit entity (regardless of whether those documents are communications with Microsoft) to the extent those documents address the operation of OpenAI's relevant business or, as a general matter, its ability to raise capital. OpenAI</p>
-----	---------	-----	--	-----	--------	-----------	--

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							confirms the phrase “relevant business” will include OpenAI’s business related to the development of GPT-based large language models. OpenAI will review and log communications between in-house and outside counsel that also include non-lawyer OpenAI employees. OpenAI will not review or log communications that are solely between OpenAI’s outside counsel and in-house counsel in response to RFP 11.
XI. Protective Order Matters							
<i>Daily News v. Microsoft Corp., Case No. 1:24-cv-3285-SHS-OTW, Dkt. 163</i>	10/16/24	<i>Daily News v. Microsoft Corp., Case No. 1:24-cv-3285-SHS-OTW, Dkt. 172</i>	OpenAI’s Opposition to the News Plaintiffs’ Disclosure of Highly Confidential Material to Dr. Baeza-Yates	OpenAI	NYT	OpenAI and NYT submitted their positions in the October 16 Status Conference agenda ahead of the October 31 conference	<p>OpenAI’s position: OpenAI objects to the disclosure of its highly confidential, internal technical and product-related documents (including source code) to Plaintiffs’ expert, Dr. Baeza-Yates, given his role as the co-founder and Chief Scientific Officer for a commercial enterprise—Theodora.AI—that, like OpenAI, is engaged in the business of “fine-tuning” large language models. OpenAI seeks protection from disclosure given the highly commercially sensitive nature of the OpenAI internal documents and source code at issue.</p> <p>Plaintiff’s position: OpenAI objects to the disclosure of its highly confidential information to Dr. Baeza-Yates because he is a co-founder and officer of Theodora.AI, an entity “engaged in the business of fine-tuning large language models.” In a case where large language models are at issue, it is unreasonable for OpenAI to oppose disclosing its highly confidential information to Plaintiffs’ technical expert simply because he works in the field. OpenAI has not argued that Theodora.AI—a six-person Chilean company primarily focused on detecting bias in Spanish Language text—is a competitor. Thus, OpenAI has not demonstrated that disclosing its highly confidential information to Dr. Baeza-Yates presents any risk of competitive harm.</p>